

STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

ZW 1949-009

Enige methoden uit de sequentieanalyse

"Actualiteiten"

J.H.B. Kemperman



29 October 1949

Voordracht door J.H.B. Kempman in de serie Actualiteiten

Enige methoden uit de sequentieanalyse.

Van een gegeven groep dingen wil men een bepaalde quantitatief meetbare grootheid z onderzoeken. Bij de n -de stap van het onderzoek wordt willekeurig uit deze groep een voorwerp getrokken, en hiervan de betreffende grootheid gemeten. Zij z_n het meetresultaat.

Men neemt aan, dat de aldus verkregen grootheid z_n een stochastische verdeling heeft, die voor iedere stap dezelfde is (deze vereenvoudiging is alleen toelaatbaar, als de gegeven groep groot genoeg is) en vervolgens stelt men een hypothese H_0 betreffende de verdelingsfunctie $F(z_n)$. Al naargelang de resultaten verkregen in de eerste n stappen, neemt men na de n -de stap één der volgende beslissingen:

1. Men aanvaardt de hypothese H_0 .
2. Men verworpt de hypothese H_0 .
3. Het onderzoek wordt voortgezet.

Het benodigde aantal stappen is dus van te voren niet vastgelegd.

De sequentietest is bepaald, wanneer men een regel heeft, die bij elke stap eenduidig aangeeft, welke beslissing moet worden genomen. Meestal beschouwt men twee functies $\psi_1(n)$ en $\psi_2(n)$ gedefinieerd voor alle gehele waarden $n \geq 0$, met $\psi_1(n) < \psi_2(n)$ en $\psi_1(0) < 0 < \psi_2(0)$. Als voor $Z_n = \sum_{j=1}^n z_j$ geldt $Z_n \geq \psi_2(n)$ (resp. $Z_n = \psi_1(n)$) dan wordt de hypothese verworpen (resp. aanvaard). Is daarentegen

$\psi_1(n) < Z_n < \psi_2(n)$, dan wordt het onderzoek voortgezet.

De sequentietest wordt vaak gebruikt om twee hypothesen H_0 en H_1 te vergelijken. Zij bijv. H_0 de hypothese, dat z een verdelingsfunctie $F_0(z)$ heeft, en H_1 de hypothese, dat z een verdelingsfunctie $F_1(z)$ heeft. Zij α (resp. β) de waarschijnlijkheid, dat H_0 wordt verworpen (resp. aanvaard), terwijl de hypothese H_0 (resp. H_1) geldt.

Opdat de test voor dit geval bruikbaar wil zijn moeten

α en β voldoende klein zijn.

Zij $E_1(n)$ de verwachtingswaarde van het benodigde aantal stappen, indien hypothese H_1 geldt. Van een bruikbare test zal men bovendien verlangen, dat $E_1(n)$ (liefst simultaan) zo klein mogelijk is, onder voorwaarde, dat $\alpha \leq \alpha_0$ en $\beta \leq \beta_0$ (waarin $\alpha_0 > 0$ en $\beta_0 > 0$ gegeven kleine grootheden voorstellen).

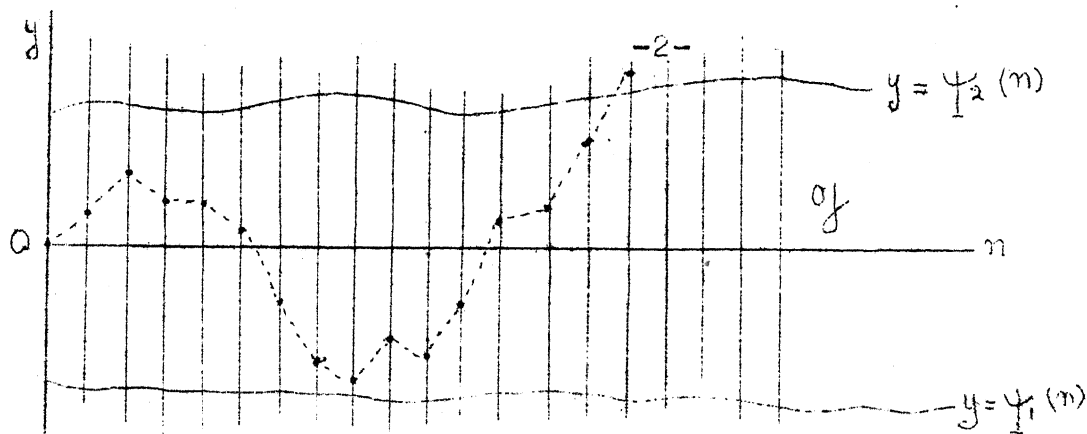


fig. 1.

Beschouw een n - y -vlak, waarin de krommen $y = \Psi_1(n)$ en $y = \Psi_2(n)$. Zij G de verzameling van de punten (n, y) met $n \geq 0$ en $\Psi_1(n) < y < \Psi_2(n)$. Elke sequentie, zoals die optreedt in een sequentietest, kan worden gerepresenteerd door de serie punten (n, Z_n) , $n=0, 1, 2, \dots$ in het n - y -vlak. Een dergelijke puntenserie noemen we een weg. Een weg gaat altijd uit van de oorsprong, beweegt zich dan voortdurend naar rechts, en eindigt in het eerste niet in G gelegen punt, de abcis van dit eindpunt komt overeen met het aantal stappen van de overeenkomstige sequentietest (we beschouwen uitsluitend punten met gehele abcis). Oneindig lange wegen verlopen geheel binnen G en omgekeerd. Een weg kunnen we verdelen in "stappen". Een stap is de overgang (n, Z_n) op $(n+1, Z_{n+1})$, d...z. we gaan een eenheid naar rechts en tegelijkertijd $Z_{n+1} - Z_n = z_n$ naar boven. Deze overgang is stochastisch met een verdelingsfunctie $P(z_n)$.

Zij $G_n(y_1)$ de waarschijnlijkheid, dat een weg de halfrechte $y \leq y_1$ passeert. Dan geldt $(\alpha_n = \Psi_1(n), \beta_n = \Psi_2(n))$

$$(1) \quad \begin{cases} G_1(y) = P(y) \\ G_{n+1}(y) = \int_{\alpha_n}^{\beta_n} P(y-\gamma) dG_n(\gamma) \end{cases},$$

daar immers een weg, die de lijn met abcis $n+1$ passeert, de lijn met abcis n gepasseerd moet zijn in een punt (n, y) met

$\alpha_n = \Psi_1(n) < y < \Psi_2(n) = \beta_n$. In principe is dus $G_n(y)$ bekend ($n=1, 2, \dots$).

De kans, dat een sequentietest leidt tot goedkeuring na precies n stappen (ofwel de kans, dat een weg eindigt in een punt (n, y) met $y \leq \Psi_1(n)$) is juist $G_n(\Psi_1(n))$ en de waarschijnlijkheid dat de steekproef ooit wordt goedgekeurd is dus $\sum_{n=1}^{\infty} G_n(\Psi_1(n))$. Analoog vinden we de kans op afkeuring en de gemiddelde lengte van een sequentie, die tot een beslissing leidt.

Het doel van de sequentieanalyse ("sequential analysis") is in de eerste plaats de berekening van de volgende grootheden,

1. De kans, dat een sequentie wordt goedgekeurd.
2. De kans, dat een sequentie wordt afgekeurd.

3. De verwachtingswaarde van de lengte van de sequentie, indien de sequentie genomen wordt uit een populatie met gegeven waarschijnlijkheidsverdeling. De boven aangegeven methode is in de praktijk meestal onbruikbaar, en we moeten dus naar andere methoden omzien.

Beschouwen we eerst het geval, dat $F(y)$ continu differentieerbaar is, met afgeleide $f(y)$, de z.g. waarschijnlijkheidsdichtheid. Als analoog $g_n(y)$ de dichtheidsfunctie is bij de verdelingsfunctie $G_n(y)$, dan volgt uit (1), dat $g_1(y) = f(y)$ en

$$(2) \quad g_{n+1}(y) = \int_{\alpha_n}^{\beta_n} f(y-\eta) g_n(\eta) d\eta,$$

als $\alpha_n = \psi_1(n)$ en $\beta_n = \psi_2(n)$. Bij de belangrijkste toepassingen geldt $\alpha_n = \alpha$, $\beta_n = \beta$ ($n=0,1,2,\dots$) en in dit geval is het handig om in te voeren

$$(3) \quad \mathcal{O}_L(u,y) = \sum_{n=0}^{\infty} u^n g_{n+1}(y).$$

Dan volgt

$$\begin{aligned} \mathcal{O}_L(u,y) &= f(y) + u \sum_{n=0}^{\infty} u^n g_{n+1}(y) \\ &= f(y) + u \sum_{n=0}^{\infty} u^n \int_{\alpha}^{\beta} f(y-\eta) g_n(\eta) d\eta \end{aligned}$$

en dus geldt (wanneer (3) gelijkmatig convergeert) de integraalvergelijking

$$(4) \quad \mathcal{O}_L(u,y) = f(y) + u \cdot \int_{\alpha}^{\beta} f(y-\eta) \mathcal{O}_L(u,\eta) d\eta$$

Als $|f(z)| \leq M$ voor $-\infty < z < +\infty$, dan heeft deze vergelijking één en slechts één oplossing voor $|u| < \frac{1}{M(\beta-\alpha)}$, welke gegeven wordt door de reeks van Neumann. Deze reeks is overigens niets anders dan de reeks (3), waarbij g_n bepaald is door de recussieformule (2) ($g_1(y) = f(y)$). In ieder geval opent (4) de mogelijkheid om de theorie van de integraalvergelijkingen toe te passen.

We beperken ons in het volgende tot het geval, dat $F(z)$ de volgende verdelingsfunctie is

$$\begin{aligned} F(z) &= 0 & \text{voor } z < -s \\ &= p & \text{" } z < +r \\ &= p+q & \text{" } z \geq r \end{aligned}$$

waarin $p+q = 1$, en waarin r en s gehele getallen voorstellen met $r+s > 0$. De grootte z neemt dus slechts de waarden $-s$ en $+r$ aan, met waarschijnlijkheid p resp. q . De meeste methodes overigens gebezigd voor dit speciale geval zijn ook bruikbaar, als z eindig vele gehele waarden z_1, z_2, \dots, z_t kan aannemen met waarschijnlijkheid

p_1 , resp. p_2, \dots , resp. p_t waarbij $p_1 + p_2 + \dots + p_t = 1$.

In de afteekening getekend in fig.1 heeft elke stap van een weg slechts twee mogelijkheden, overeenkomend met de overgangen $(n,y) \rightarrow (n+1, y-s)$ en $(n,y) \rightarrow (n+1, y+r)$, waarvan de kansen zich verhouden als p tot q .

Opdat een punt (n,y) bereikt kan worden door een weg, is nodig, dat er niet negatieve gehele getallen j en k bestaan met

$$(5) \quad j + k = n \quad ; \quad kr - js = s.$$

Bij gegeven n zijn er slechts eindig veel waarden y , waarvoor dit het geval is.

Onder een toelaatbare stap in het n - y -vlak verstaan we de overgang $(n,y) \rightarrow (n+1, y-s)$ of $(n,y) \rightarrow (n+1, y+r)$ met het beginpunt (n,y) in \mathcal{G} . Als P en Q roosterpunten zijn (d.w.z. punten in het rechterhalfvlak met gehele coördinaten), dan zullen we onder een "toelaatbare baan van P naar Q " verstaan een aansluitende reeks toelaatbare stappen met als beginpunt P en als eindpunt Q .

Aan elk roosterpunt (n,y) van \mathcal{G} voegen we nu een geheel getal $\alpha(n,y)$ toe, dat het aantal toelaatbare banen aangeeft van 0 naar (n,y) . Voor een punt (n,y) buiten \mathcal{G} stellen we $\alpha(n,y) = 0$. Aan de hand van fig.1 zien we in, dat voor een van 0 verschillend roosterpunt (n,y) geldt

$$(6) \quad \alpha(n,y) = \alpha(n-1, y+s) + \alpha(n-1, y-r).$$

Mgens $\alpha(0,0) = 1$ kan men in principe voor een gegeven roosterpunt (n,y) de grootheid $\alpha(n,y)$ berekenen. Het zal blijken, dat na bepaling van $\alpha(n,y)$ voor elk punt (n,y) van \mathcal{G} alle gezochte functies, zoals b.v. de kans op afkeuring (resp. aanvaarding), de verwachtingswaarde van de steekproeflengte (=weglengte) direct kunnen worden neergeschreven.

Laten r' en s' gehele getallen zijn met $r'+s' \neq 0$ en beschouw nu de punttransformatie

$$(7) \quad \begin{cases} n' = n \\ \frac{y'}{r'+s'} - \frac{y}{r+s} = \frac{1}{2} n \left(\frac{r'-s'}{r'+s'} - \frac{r-s}{r+s} \right) \end{cases}$$

De overgang $(n,y) \rightarrow (n+1, y+r)$ resp. $(n,y) \rightarrow (n+1, y-s)$ in het n - y -vlak correspondeert nu met een overgang $(n',y') \rightarrow (n'+1, y'+r')$ resp. $(n',y') \rightarrow (n'+1, y'-s')$ en omgekeerd. De randkrommen $y = \psi_i(n)$ ($n=1,2$) in het n - y -vlak corresponderen met krommen

$$(8) \quad y' = \psi_i^*(n) = \frac{r'+s'}{r+s} \psi_i(n') + \frac{1}{2} n' \left[(r'-s') - (r-s) \frac{r'+s'}{r+s} \right]$$

die het beeld \mathcal{G}' van \mathcal{G} begrenzen. Toelaatbare banen in het n - y -vlak

f van toelaatbare banen in het r' - y' -vlak lopend van 0 naar (n', y') worden eeneenduidig afgebeeld op toelaatbare banen in het n' - y' -vlak (samengesteld uit stappen ter lengte r' resp. s' ; alleen punten van j' mogen worden gepasseerd). Voor een punt (n', y') van j' is dus het aantal $\beta(n', y')$ precies gelijk aan het aantal toelaatbare banen $\alpha(n, y)$, die in het n - y -vlak van 0 lopen naar het origineel (n, y) van (n', y') .

Wanneer voor de begrenzingsen $y = \psi_i(n)$ en toegelaten stappen ter lengte r resp. s de functie $\alpha(n, y)$ bekend is, dan is voor de begrenzingsen $y = \psi_i^*(n)$ en stappen ter lengte r' resp. s' de corresponderende functie $\beta(n, y)$ dus eveneens bekend (en omgekeerd).

Neem aan, dat de grootte z , die gemeten wordt bij het sequentieonderzoek, de gehele waarden r en s ($r+s \neq 0$) aanneemt met waarschijnlijkheid q resp. p ($p+q=1$), en laten de randkrommen bepaald zijn door

$\psi_1(n) = a_1 + \mu n$; $\psi_2(n) = a_2 + \mu n$,
waarin $a > 0$ en $a_2 < 0$ reële constanten zijn en waarin μ rationaal is. Wil men voor alle mogelijke combinaties r, s, a_1, a_2 , dit sequentieprobleem oplossen, dan kan men zich beperken tot het geval $\mu = 0$.

Indien n.l. $\mu \neq 0$, dan bestaan gehele getallen r' en s' (met $r'+s' \neq 0$), en constanten a'_1 en a'_2 zo, dat

$$(9) \quad a'_i = \frac{r'+s'}{r+s} (a_i + \mu n') + \frac{1}{2} n' \left[(r'-s') - (r-s) \frac{r'+s'}{r+s} \right]$$

($i=1,2$; $n=0,1,2,\dots$), daar voor r' en s' met

$$\frac{r'-s'}{r'+s'} = \frac{r-s-2}{r+s}$$

het rechterlid van (9) een constante is.

We noemen de begrenzing van j ($y = \psi_1(n)$ resp. $y = \psi_2(n)$) „op den duur periodiek”, als een rationaal getal μ en een natuurlijk getal ω bestaat met

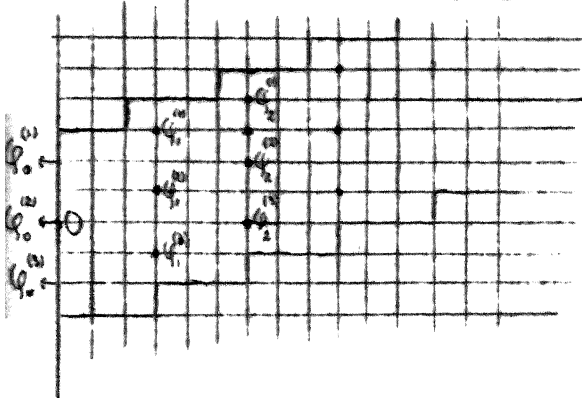
(10) $\psi_1(n+\omega) = \mu \omega + \psi_1(n)$; $\psi_2(n+\omega) = \mu \omega + \psi_2(n)$,
voor n voldoende groot. In dit geval bestaat er een algemene methode. Voor de berekening is het 't gemakkelijkst om het probleem te transformeren op het geval $r=s=1$. Past men transformatie (7) toe met $r'=s'=1$, dan gaan de formules (7) en (8) over in

$$(11) \quad y' = \frac{2}{r+s} y + \frac{s-r}{s+r} n'$$

en

$$(12) \quad \psi_i^* = \frac{2}{r+s} + \frac{s-r}{s+r} n'$$

Uit (10) en (12) volgt, dat het beeld σ_j^1 van σ_j weer periodieke begrenzingsen heeft. We mogen ons dus beperken tot het geval $r=s=1$. Aan het volgende voorbeeld herkent men gemakkelijk de algemene methode (welke ook zonder de transformatie (11) toepasbaar is op het geval van willekeurige paren gehele waarden r en s met $r+s \neq 0$).



Zij $\psi_1(n) = -3 + \frac{1}{3}n$, $\psi_2(n) = \frac{4}{3} + \frac{1}{3}n$; als deze functies de randkromme bepalen, dan kunnen we de rand evengoed representeren door de getekende trapvormige krommen, daar dit de verzameling van de roosterpunten tussen de twee randkrommen niet wijzigt.

Stel nu

$$Q_1^{(1)} = (31, 1+2) \text{ en } \alpha(31, 31+2) = b_1^{(1)}$$

$$Q_1^{(2)} = (31, 1) \text{ " } \alpha(31, 1) = b_1^{(2)}$$

$$Q_1^{(3)} = (31, 1-2) \text{ " } \alpha(31, 1-2) = b_1^{(3)}$$

Hierin is dus $b_{l+1}^{(m)}$ het aantal toegelaten banen van 0 naar $Q_{l+1}^{(m)}$. Elke dusdanige baan passeert precies één der punten $Q_1^{(1)}$, $Q_1^{(2)}$, $Q_1^{(3)}$. Zij γ_{ik} het aantal toegelaten banen van $Q_1^{(k)}$ naar $Q_{l+1}^{(i)}$. Essentieel is, dat deze grootheden γ_{ik} onafhankelijk zijn van l , zoals direct volgt uit de periodiciteit van de rand. Het aantal banen vanaf 0, dat in $Q_{l+1}^{(i)}$ aankomt en het punt $Q_1^{(1)}$ gepasseerd is, is nu $\gamma_{i,1} b_1^{(1)}$. Dus volgt

$$(13) \quad b_{l+1}^{(i)} = \sum_{m=1}^3 \gamma_{im} b_1^{(m)} \quad (i=1,2,3).$$

Vorm nu

$$(14) \quad g_i(t) = \sum_{l=0}^{\infty} b_l^{(i)} t^l.$$

Dan volgt

$$g_i(t) = b_0^{(i)} + t \sum_{l=0}^{\infty} b_{l+1}^{(i)} t^l = b_0^{(i)} + t \sum_{l=0}^{\infty} t^l \sum_{m=1}^3 \gamma_{im} b_l^{(m)}$$

en dus

$$(15) \quad g_i(t) = b_0^{(i)} + t \sum_{m=1}^3 \gamma_{i,m} g_m(t) \quad (i=1,2,3)$$

Hierin is $b_0^{(1)}=0$, $b_0^{(2)}=1$, $b_0^{(3)}=0$. Na bepaling van $\gamma_{i,m}$ kan men uit (15) dus $g_m(t)$ berekenen¹⁾. Nu is $\gamma_{i,k}$ bij definitie gelijk aan het

1) De determinant $|\delta_{ij}^m - t \gamma_{im}|$ van het stelsel vergelijkingen is voor voldoende klein t zeker van nul verschillend.

aantal toegelaten banen van $Q_0^{(k)}$ naar $Q_1^{(i)}$. Als we in het bijzonder $\chi_{i,2}$ willen bepalen, dan geldt wegens $Q_0^{(2)} = 0$, dat

$$\chi_{1,2} = \alpha(3,3) ; \chi_{2,2} = \alpha(3,1) ; \chi_{2,3} = \alpha(3, .1),$$

waarin $\alpha(n,y)$ met

$$\alpha(n+1, y) = \alpha(n, y-1) + \alpha(n, y+1) \text{ en } \alpha(0,0)=1 \text{ direct uit}$$

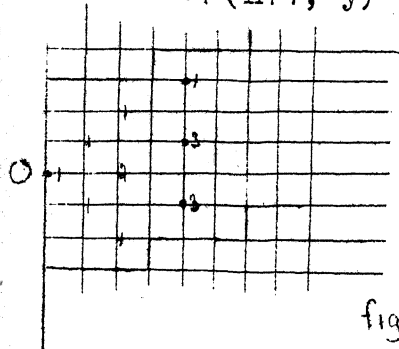


fig.3

de figuur kan worden bepaald. We vinden met behulp van nevenstaande figuur, dat $\chi_{1,2}=1$, $\chi_{2,2}=3$, $\chi_{3,2}=3$. Op geheel analoge wijze vinden we de andere transformatiecoëfficiënten $\chi_{i,k}$. Het resultaat is

$$(\chi_{i,k}) = \begin{pmatrix} 1 & 1 & 0 \\ 2 & 3 & 1 \\ 1 & 3 & 2 \end{pmatrix}.$$

De oplossing van (15) wordt nu

$$(16) \quad \begin{cases} g_1(t) = \frac{t - 2t^2}{1-6t+6t^2} & = -\frac{1}{3} + \frac{1/3-t}{1-6t+6t^2} \\ g_2(t) = \frac{1-3t+2t^2}{1-6t+6t^2} & = +\frac{1}{3} + \frac{2/3-tt}{1-6t+6t^2} \\ g_3(t) = \frac{3t-2t^2}{1-6t+6t^2} & = -\frac{1}{3} + \frac{1/3+t}{1-6t+6t^2} \end{cases}$$

Stel nu

$$\sum_{n=0}^{\infty} c_n t^n = \frac{1}{1-6t+6t^2} = \frac{1}{[1-(3+\sqrt{3})t][1-(3-\sqrt{3})t]} =$$

$$= \frac{t^{-1}}{2\sqrt{3}} \left[\frac{1}{1-(3+\sqrt{3})t} - \frac{1}{1-(3-\sqrt{3})t} \right]$$

dan vinden we

$$(17) \quad c_n = \frac{1}{6\sqrt{3}} \left[(3+\sqrt{3})^{n+1} - (3-\sqrt{3})^{n+1} \right] \quad (n=0,1,2,\dots)$$

en wegens (13) en (16) volgt dan voor $l > 0$

$$b_1^{(1)} = \frac{1}{3} c_1 - c_{1-1}$$

$$b_1^{(2)} = \frac{2}{3} c_1 - c_{1-1}$$

$$b_1^{(3)} = \frac{1}{3} c_1 + c_{1-1}.$$

Onder een randpunt P van \mathcal{O} verstaan we altijd een punt buiten \mathcal{O} , dat bereikt kan worden door een toegelaten weg beginnende in 0 (en die dus behalve P geen roosterpunten buiten \mathcal{O} bevat). Als (n,y) randpunt is dan duiden we met $R(n,y)$ het aantal toegelaten wegen aan lopende van 0 naar (n,y) . Zij l geheel ≥ 0

Een toegelaten weg vanaf 0 , die eindigt in een randpunt (n,y) met $3l < n \leq 3l+3$ is noodzakelijk één der punten $Q_1^{(i)}$ ($i=1,2,3$) gepasseerd,

en moet dus noodzakelijk in één der punten U_1 , V_1 of W_1 (zie fig.4) eindigen. Tevens volgt uit de figuur, dat in het punt $U_1 = (3l+1, 1+3)$ evenveel wegen aankomen als in het punt $Q_1^{(1)}$. Dus volgt $R(3l+1, 1+3) = b_1^{(1)} = \frac{1}{3} c_1 - c_{1-1}$. Analoog vindt men voor V_1 en W_1 , dat $R(3l+1, 1-3) = b_1^{(3)} = \frac{1}{3} c_1 + c_{1-1}$ en $R(3l+3, 1-3) = b_1^{(2)} + b_1^{(3)} = c_1$.

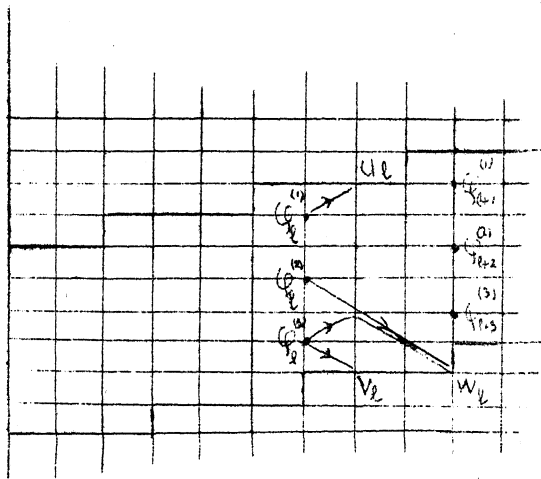


fig.4

Van elk randpunt P kennen we nu aantal toegelaten wegen van O naar P . We willen nu waarschijnlijkheid weten, dat een van O uitgaande weg in P eindigt, als bij elke stap de overgang $(n,y) \rightarrow (n+1, y-1)$ een kans p heeft en de overgang $(n,y) \rightarrow (n+1, y+1)$ een kans q , ($p+q=1$). Een weg, die het

randpunt (n,y) bereikt na j maal een stap $(n',y') \rightarrow (n'+1, y'-1)$ en k maal een stap $(n',y') \rightarrow (n'+1, y'+1)$ gedaan te hebben heeft dus een waarschijnlijkheid $p^j q^k$. Wegens

$$j + k = n, \quad -j + k = y$$

ofwel

$$j = \frac{n-y}{2}, \quad k = \frac{n+y}{2}$$

is deze waarschijnlijkheid dus gelijk aan $p^{\frac{n-y}{2}} q^{\frac{n+y}{2}}$. De waarschijnlijkheid, dat een weg in een randpunt (n,y) eindigt wordt dus $R(n,y) = p^{\frac{n-y}{2}} q^{\frac{n+y}{2}}$. Dit betekent dus ($l > 0$)

voor $U_1 = (3l+1, 1+3)$ een waarschijnlijkheid $(\frac{1}{3}c_1 - c_{1-1})p^{1-1}q^{2l+2}$

$$(18) \quad " \quad V_1 = (3l+1, 1-3) \quad " \quad " \quad (\frac{1}{3}c_1 - c_{1-1})p^{1+2}q^{2l-1}$$

$$" \quad W_1 = (3l+3, 1-3) \quad " \quad " \quad c_1 p^{1+3}q^{2l}$$

De in het randpunt afgebroken weg bestaat in deze gevallen uit $3l+1$ resp. $3l+1$ resp. $3l+3$ stappen. Opgemerkt moet nog worden dat U_0 en V_0 , in boven aangegeven zin, geen randpunten zijn. De waarschijnlijkheid, dat een weg in W_0 eindigt is p^3 , zoals ook uit (18) zou volgen ($c_0=1$)

Alle functies, die in de sequentieanalyse gezocht worden kunnen (in ons voorbeeld) nu meteen worden neergeschreven. Bijvoorbeeld de kans, dat een weg wordt afgebroken aan de bovenrand

(d.w.z. de kans dat een sequentie wordt afgekeurd) is nu gelijk aan

$$\sum_{l=1}^{\infty} \left(\frac{1}{3} c_l - c_{l-1} \right) p^{l-1} q^{2l+2} = \frac{q^2}{3p^2} \sum_{l=1}^{\infty} c_l (pq^2)^{l-1} - q^4 \sum_{l=0}^{\infty} c_l (pq^2)^l =$$

$$= \frac{q^2}{3p} (1 - 3pq^2) \cdot \frac{1}{1 - 6pq + 6p^2q^2 - \frac{q^2}{3p}},$$

waarbij we gebruik maakten van de definitie $\sum_{\ell=0}^{\infty} c_\ell t^\ell = \frac{1}{1 - 6t + 6t^2}$

Naar ik meen, is aan dit voorbeeld voldoende duidelijk geworden hoe men het algemene geval van een periodieke begrenzing van γ kan behandelen. De kans op goedkeuring en afkeuring van een sequentie (volgens een lemma van A. Wald zijn deze samen 1), de gemiddelde lengte van een sequentie, en in het algemeen het gemiddelde van n^k (k geheel ≥ 0 en overigens willekeurig), zijn alle rationale functies van p en q in het geval van een periodieke begrenzing.

Opm.: Als in (13) de coëfficiënten γ_{im} bepaald zijn, dan kan men in plaats van de genererende functies (14) in te voeren ook vragen naar de eigenwaarden en eigenvectoren van de transformatie (13). Neem aan, dat deze bepaald zijn: zij λ_m een eigenwaarde met genormeerde eigenvector $(x_m^{(1)}, x_m^{(2)}, x_m^{(3)})$, $m=1,2,3$. Zijn de eigenwaarden alle verschillend dan zijn de eigenvectoren orthogonaal. De vector $(b_1^{(1)}, b_1^{(2)}, b_1^{(3)})$ is uit de vector $(b_0^{(1)}, b_0^{(2)}, b_0^{(3)}) = (0,1,0)$ ontstaan door de transformatie (13) 1 maal achtereenvolgens toe te passen. Uit

$$(b_0^{(1)}, b_0^{(2)}, b_0^{(3)}) = \sum_{m=1}^3 x_m^{(2)} \cdot (x_m^{(1)}, x_m^{(2)}, x_m^{(3)})$$

volgt nu

$$(b_1^{(1)}, b_1^{(2)}, b_1^{(3)}) = \sum_{m=1}^3 \lambda_m^{-1} x_m^{(2)} (x_m^{(1)}, x_m^{(2)}, x_m^{(3)}).$$

De absoluut grootste onder de eigenwaarden λ_m ($m=1, \dots$) bepaalt dus het asymptotisch gedrag van $b_1^{(i)}$ en in het bijzonder ook de convergentiestraal van (14), aangenomen dat voor deze grootste eigenwaarde λ_m geldt $x_m^{(2)} \neq 0$.

Natuurlijk verschilt de methode, die eigenwaarden en eigenvectoren bezigt, niet essentieel van de methode, die de genererende functies (14) gebruikt. De determinant van het stelsel vergelijkingen (15) is n.l. $|\delta_i^m - t \gamma_{i,m}| = \Delta(t)$, en de nulpunten van $\Delta(t)$ zijn dus juist de **reciproken** van de **eigenwaarden** λ_m , en deze nulpunten λ_m^{-1} spelen een fundamentele rol in de ontwikkeling van de uit (15) verkregen functies $g_i(t)$.